Defense Technical Information Center

Compilation Part Notice

ADP012335

TITLE: Agent-Based Modelling for Environmental Coalition Formation

DISTRIBUTION: Approved for public release, distribution unlimited

Availability: Hard copy only.

This paper is part of the following report:

TITLE: KSCO 2002: Second International Conference on Knowledge Systems for Coalition Operations

To order the complete compilation report, use: ADA402533

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP012330 thru ADP012354

Agent-Based Modelling for Environmental Coalition Formation

Jim Doran

Department of Computer Science
University of Essex
Colchester, UK, CO4 3SL
doraj@essex.ac.uk

Abstract. Planned intervention to achieve stakeholder cooperation and coalition is essential for successful environmental management. Agent-based modelling on a computer has the potential to build a practical theory of intervention in this and related contexts. Potentially we can compare real intervention strategies with those an agent-based model suggests and hence obtain new insights and guidelines of practical value. But the technical problems of model building for this purpose are formidable. We explain and discuss these problems by reference to an example model specification framework, and seek ways forward. Insights obtained may be generalised to coalition formation in general.

1 Introduction

Based on multi-agent systems (MAS) theory (Weiss, 1999), computer and agent-based modelling of social and organisational systems (Doran, 1997, 2001a) is becoming of practical value in a range of application domains (Moss and Davidsson, 2001) including the military (Tessier et al., 2000), the environmental (Bousquet et al., 1999) and the social (Gilbert, 2000).

Here we take the view that a multi-agent system is an interacting collection of agents sharing a common (possibly simulated) environment, where an agent may loosely be viewed as an "object" in the software engineering sense that possesses a degree of autonomy and a modicum of cognitive ability. This indicates the relevance of artificial intelligence theory and practice (Russell and Norvig, 1995).

Cooperation is a key topic in MAS (Doran et al., 1997). Most agent work on cooperation concerns how to design cooperation into a MAS, or how to model existing cooperation, rather than how to achieve it in a pre-existing non-cooperating set of agents. But *achieving cooperation* in a pre-existing situation is very often the real-world problem. We view real-world *coalitions* as involving the mutually agreed temporary cooperation of large organisations without loss of organisational identity or rights. Often the word "coalition" has international military or national political connotations.[1] However, Keohane and Ostrom (1995) have demonstrated the close relationship between cooperation for the solution of environmental problems, and more general international cooperation.

2 An Environmental Problem: Integrated Watershed Management

Integrated watershed management is the task of organising the activities and requirements in a river basin to achieve multiple and conflicting goals (Abu-Zeid and Biswas, 1996; Westervelt, 2000). Stakeholder cooperation is essential. Typically there are conflicting requirements to be balanced of:

- water supply (domestic, agricultural, industrial uses)
- pollution control
- fisheries management
- flood control
- hydropower production
- navigation and wetlands management
- recreation provision

Always there will be many stakeholders associated with different activities in the basin, all with their own objectives and agendas. Conflicts of interest are inevitable. A good example of this is the Fraser River basin in British

---

[1] Compare the connotations of the word "consortium".

Columbia (Healy, 1999; Doran, 2001). Large-scale *river engineering* projects can involve even wider issues, but are beyond the scope of this short paper.[2]

3 Interventions and Models of Intervention

Integrated watershed management, and similar ecosystem management problems, typically involve *intervention*. That is, some person, some group or some organisation, has the task of intervening in the ecosystem in order to bring about desirable change, often using the notion of a search for *sustainability*. The intervener may be, for example, a branch of the UN, an NGO, an academic research team or even a lone doctoral student. The practice of intervention is so much a part of the ecosystem management task that, in our view, it is unrealistic to ignore it for modelling purposes. The intervention history of the Fraser River Basic is a revealing example of just what issues can arise in intervention, and what can go wrong (Dorcey, 1997; Marshall, 1998; Doran, 2001).

It is evident that there can be a range of *intervention strategies*. A number of these have been discussed in Doran (2001). Here we are particularly interested in a two-stage intervention process, in which intervention first seeks to build an effective coalition and only then to set that coalition into action on the actual management task.

Symbolically we may write the intervention task as:

$$INTERVENTION \rightarrow (MAS + ENVSYS)$$

or recognising, that coalition formation may be part of the intervention process, as:

$$INTERVENTION \rightarrow COALITION \rightarrow (MAS + ENVSYS)$$

We would like to model *all* of this intervention process on a computer in order to explore possible intervention strategies with the minimum of habitual and cultural pre-conceptions.

3.1 Essentials of a Typical Environmental Resource Management Problem

We assume that environmental "harvesting"[3] requires:

- distributed action coordinated in space and time.

Furthermore actors (individual or organisational) must show restraint if they are to achieve, as we shall require:

- collective long-term survival (i.e. *sustainability*)
- the *protection* of specified environmental components
- some kind of *equity*[4] between actors

The central difficulty is that human beings tend to be individually, collectively and organisationally "greedy" and with bounded rationality. In particular, we tend to think short-term. Any potentially informative model must capture these characteristics. Compare "common pool resource" (CPR) problems of which this formulation may be seen as generalisation (Hardin, 1968; Ostrom, 1990, 1995).

3.2 A Research Plan

For clarity and focus, we foreground the following five-stage computer-based research plan:

1. Formulate a representative ENVSYS in mathematical/computational terms. The ENVSYS must reward distributed coordination and embody the sustainability, equity and protection problems identified above. Examine its long-term dynamics.

---

[2] The Three Gorges Project on the Yangtze, for example, involves further issues such as massive population movement and destruction of archaeological sites – and for this and other reasons has become highly politically charged.
[3] The work "harvesting" is here used in an extended sense to cover the collective exploitation of natural resources.
[4] Not all would agree that the last of these requirements, the equity requirement, should be included in a general definition of natural resource management.

2. Generate a sample of MAS connected to the ENVSYS. They should be neither incoherent nor successfully achieving sustainability, protection and equity over the chosen time span, that is, the generated MAS should function but fail to solve the problems.

3. Try to interpret the generated sample MAS in first abstract then human/social terms. This will probably include recognition of different types of MAS.

4. Search the space of all possible interventions to find those that are most successful for MAS of each type, where success refers to a high degree of maintenance of harvest, without depletion of protected environmental components, and with equal distribution of harvest over the set of agents.

5. Interpret the interventions found in both abstract and human/social terms

Throughout the execution of such a research plan it is essential not to confuse two distinct domains of investigation:

- Intervention to achieve cooperation in a human social system with initially conflicting stakeholders

- Intervention to achieve cooperation in an abstract MAS on a computer with initially conflicting agents

It is the latter computational domain to which the research plan directly refers. The central questions are whether effective intervention strategies can be identified in the computational domain, and then whether or not these identified intervention strategies have relevance to the real world domain.

4 A Framework for a Model

To proceed we need a precise and programmable specification of a MAS+ENVSYS and of possible interventions upon it that is sufficiently realistic for conclusions drawn from it to be reliable. In spite of all the advances made in agent technology and artificial intelligence over the past half century, this is difficult to achieve. The following framework should therefore be regarded as, at best, pointing the way ahead.

4.1 Two Basic Assumptions

We work from two basic assumptions, both controversial. The first is:

All social phenomena can in principle be captured within a computer-based model

This is analogous to the strong AI assumption that all aspects of intelligence can be captured within a computer-based model. It lies at the heart of multi-agent-based social modelling, but certainly not all social scientists or practitioners of agent-based social modelling would subscribe to it. Its significance here is that it encourages us to be optimistic that the model we want can in principle be found.

The second assumption is:

The social is emergent from the individual and the neural, and should be modelled accordingly

If anything this is even more controversial, for it is strongly reductionist and therefore unfashionable. Its significance here is that it suggests that to design and build an explicitly high-level social model is to omit its most important property, *emergence* (see, for example, Conte and Gilbert, 1995, pp 8-12). Rather the objective must be to explore the space of *low-level* models, seeking those that display high-level emergent phenomena and structures. Thus the specification that follows delineates a *class* of models rather than a specific model. Indeed, the aim is not to design a model ourselves, but rather to discover what models are *possible*, employing in effect a process of intelligently designed and efficient "generate and test".

4.2 ENVSYS

An ENVSYS is structured as a set of Boolean, integer or real-valued variables inter-related by recurrence relations of the general form

$$x_n(t+1) = f(x_1(t) \ldots \ldots x_q(t))$$

where t refers to time and the subscripts index variables.

It is *not* intended that the ENVSYS be a model of a particular real-world environmental system. Rather *the recurrence relations, together with the "actions" available to the agents (see later) and the agents' "localities" (see later), should be chosen to provide the required resource management problem characteristics, that is, the need for distributed and coordinated harvesting together with difficulty in achieving sustainability, protection and equity* (see section 3.1). Distributed and coordinated harvesting may be a matter of a specified pattern of actions upon a particular set of variables (actions and variables distributed in time as well over localities) having a disproportionate and "beneficial" impact upon key harvestable variables. Motivating real-world instances range from large-scale irrigation systems and specialised artefact production to simple group cooperation activities such as ditch digging and tree felling. Problems of sustainability (and protection) may be posed by so choosing the ENVSYS relations that harvesting beyond a certain amount results in the harvestable (or protected) variables being driven beyond acceptable limits or permanently set to zero. Equity is naturally expressed as the requirement that all agents harvest to roughly the same degree.

The ENVSYS may be formulated in many ways. For example, the recurrence relations may form something akin to a classic systems dynamics model (see Westervelt, 2000). Alternatively, the ENVSYS may be more in the tradition of "Artificial Life" studies with a spatial interpretation that has agents moving and harvesting localised resources on a plane (e.g. Epstein and Axtell, 1995).

4.3 MAS Agents

Agents must harvest at a minimum total rate or they are deleted.

Each agent is structured as a set of tokens, the contents of its working memory (WM), together with condition-action rules that execute upon and manipulate the working memory and which observe and manipulate the agent's external context.

Tokens

EITHER a simple token

a (bounded) string of letters, possibly prefixed by not (the negation character)

OR a variable-value token

a pair: a (bounded) string of letters, and a value

Rules

A pair:

a (bounded) set of tokens and a (bounded) set of actions

where an action is of one of the following types:

Harvest -- deplete a specified ENVSYS variable by a specified amount

Set -- set a specified ENVSYS variable to a specified value

Read -- read the value of a specified ENVSYS variable and deposit a corresponding variable value token in the WM

Deposit -- deposit a specified token in (own) WM

Send -- deposit a specified token in WM of another specified agent

Locality

Each agent has its own locality, which is fixed in time, that is, each agent can set, read and harvest a specified subset of the variables – its "local" variables.

4.4 Agent Processing

Rule Firing {

```
        Find all rules whose LHS match in the current WM where a match
        requires that every LHS token occurs in the current WM

        Select a matched rule at random

        Execute the selected rule's action(s)
    }
```

Token Reconciliation

We say that two tokens *contradict* if they differ only in the negation character.

It is assumed that the initial contents of the WM are contradiction free. If a token is introduced (by an internal or external rule firing or an intervention) that contradicts an existing token (i.e. differs from it only in the negation character) then the pre-existing token is deleted from the WM. This conflict removal procedure is very simplistic and certainly not, of course, logically complete.

Rule Set Reconciliation

We say that two rules *contradict* if their conditions are identical but their actions differ.

It is assumed that the initial rules set is contradiction free. If a rule is introduced into the rule set (by intervention) that contradicts an existing rule, then the pre-existing rule is deleted. Again, this conflict removal procedure is not logically complete.

4.5 Intervention

An *intervention element* is the deposition of one token or one rule into a particular agent's working memory at time t. An *intervention* is a set of N intervention elements. The impact of an intervention element is determined by the reconciliation procedure.

4.6 Processing MAS+ENVSYS + interventions

Initialise MAS+ENVSYS at random and set clock to zero

```
    Repeat
    {
            Advance clock (t)

            Activate each agent once
                 (in a varying random order)

            Pass any inter-agent messages

            Apply any interventions at this time

            Reconcile each agent's tokens and rules

            Update the ENVSYS

            Collect statistics
    }
    until time limit reached
```

This semi-formal specification is, in artificial intelligence terms, very simple. "Filled in" with rules and initial token sets for the agents' working memories, it is clearly programmable (in, for example, C++) and model instances can therefore certainly be "run" and experimented with. However, the combination of tokens and rules is computationally sufficiently powerful that complex cognitive processes such as learning and planning are certainly possible. It is not easy to anticipate in any detail what specific types of dynamics will occur within an agent or a

MAS in particular circumstances.[5] Nevertheless some aspects of the behaviour of this type of model are foreseeable, as we shall discuss in the next section.

5 Major Characteristics of the Model

We now turn to the foreseeable characteristics of this model, and the technical difficulties that any attempt to use it will encounter.

5.1 Properties and Problems

It is important to appreciate that complex cognitive processes, for example the use of internal representations, goal setting, plan formation and execution, and learning, are potentially present in an agent's working memory dynamics even though agents are "merely" rule based. This follows from the fact that the contents of an agent's working memory both determine and are modifiable by the rules that "fire". That does not mean, of course, that agents with cognitive processes are easily generated nor, less obviously, that it is easy to recognise them when they are. Indeed, just how cognitive processes can be recognised in practice in such a context is an interesting and far from trivial question.

The behaviour of any particular instance of the model that meets the specified requirements (a *solution model instance*) is primarily determined by the rule sets within the agents. To serve our purposes, these rule sets must be such that the MAS, without intervention, has the specified properties with respect to the ENVSYS, notably that it does successfully "harvest" resources, but not so that it is immediately sustainable, equitable and protective. But the probability that an arbitrary or randomly generated MAS will function in this way, or even function coherently, is very small indeed. There is therefore a significant combinatorial problem merely to find functioning and effective MAS. Some form of "hill-climbing" algorithm or evolutionary algorithm[6] could be used, at least on the micro-scale. Just how complex are the effective MAS that could be found in this way is an open question. Of course, one could set out explicitly to *design* an effective MAS (a kind of programming exercise) but this would be to pre-determine what we wish to discover, and it encounters head-on the difficulty that our ability to program the needed artificial intelligence capabilities is limited. A compromise might be to design some basic structures and capabilities into the model's agents, perhaps sufficient for their minimal survival by purely uncoordinated action in the ENVSYS, and to leave the rest to some form of heuristic or evolutionary search.

Once discovered, effective MAS may or may not display (emergent) collections of agents that may reasonably be labelled "organisations" (compare Prietula et al., 1998). They may or may not display centralised decision-making and/or collective planning. Agents (and agent organisations) will typically be heterogeneous, perhaps in a patterned way and, as just suggested, may or may not incorporate cognitive processes. All discovered MAS are likely to be "noisy" in the sense that their rules and working memory contents will often include much that is inessential to their required functioning.

Recall that the purpose of generating MAS that can successfully interact with the ENVSYS is precisely to *discover* what form such MAS can take (rather then prejudge that issue) and to then take the next step to consider intervention.

5.2 Interventions and Intervention Strategies

Organised patterns of intervention (*intervention strategies*) may be *discovered* to be structured in various ways, and they may either prompt a successful pattern of action, or may prompt a social structure (e.g. a coalition) which will itself achieve the required pattern of action, or may prompt something even more complex.

Assuming a fixed instance of a MAS+ENVSYS, optimal interventions can be defined and (in principle) determined without addressing the issue of the intervener's knowledge of MAS+ENVSYS. However, this issue cannot be avoided if the requirement is changed to that of finding a *decision procedure* that gives an effective intervention. Such a decision procedure would be a function of the intervener's knowledge of the MAS+ENVSYS.

5.3 Translation to and from the Model

To make practical use of a solution model instance requires that we are clear about the structural relationship between the two domains of intervention strategy. For example, what corresponds in the abstract model to

---

[5] Compare Turing Machines (Turing, 1937), and also the well-known Agent0 agent-oriented programming language (Shoham, 1993).
[6] We are here using the phrase "evolutionary algorithm" in a technical sense. There is no question of modelling human evolution.

centralisation and decentralisation? social capital? organised conflict? a coalition? And how may these specifically be achieved by intervention? Here we focus briefly on coalitions.

5.3.1 Coalitions

Assuming the model specification of section 4, and given our initial attempted definition of a coalition as "involving the mutually agreed temporary cooperation of large organisations without loss of organisational identity or rights", what form would a coalition take in such a model, under what circumstances might intervention lead to the formation of a coalition, and when might that coalition be effective?

It seems reasonable to suggest that we are looking for a set of agents that are in some sense "leaders of" organisations and that further, for a significant period of time:

- have a pattern of inter-communication amongst them, and
- display some degree of shared goals, and
- display a degree of coordinated action.

It follows that the recognition of coalitions in a MAS rests upon the recognition of lower-level phenomena such as goals, communication and coordinated action. But more is needed: specifically a precise account of just what is involved in the formation, action and dispersal of coalitions. A possible basis is the formal account of the various stages of a group cooperation and action process provided by Wooldridge and Jennings (1999). Although their account is formulated in terms of a quantified multi-modal logic, and at first sight seems too abstract to be helpful here, in fact it does go at least part way to providing the kind of precise recognition procedure required. If a recognition procedure can be established, it then becomes feasible to address the ways in which different interventions strategies impact upon the MAS+ENVSYS combination, and to identify those classes of intervention strategy that lead to effective coalition formation.

6 Discussion

It may be argued that a study of this type can have very little practical value, since (i) only the simplest solution model instances can be found however sophisticated the combinatorial search procedure deployed, and these models will therefore be unrepresentative, and (ii) there are deeper reasons, in any case, why such models can *never* be relevant to real human social situations.

Point (i) seems unduly pessimistic. The success of techniques for finding solutions to complex problems by evolutionary and other heuristic techniques is well known. To assume that they will be useless in this context is surely unjustified. Furthermore, the structure of the problem, involving specific and well-defined requirements that must be met, means that the search for solution models is through a space that is in fact quite tightly defined. Coupled with ever increasing available computer power, it is at least feasible that interesting discoveries may be made.

The second objection (ii) is essentially a "philosophical" one based upon a perception that there is something intrinsically different about human society compared with an artificial agent society. In particular, it is a perception that human and agent societies must differ in how they collectively address resource acquisition and distribution tasks. This perception runs counter to our initial assumption that all social phenomena can be captured within an agent-based model. More importantly, it also runs counter to much current research that assumes and demonstrates that there is indeed a fruitful basis for the exchange of ideas about the two types of society. Is it really the case that, say, groups of robots and groups of humans faced with the same foraging task will never deploy similar strategies?

7 Conclusions

We have suggested how social intervention strategies can be discovered and classified in the abstract by generating and exploring a "space" of relevant agent-based models. The objective is to match discovered abstract strategies to those in actual "everyday" use, and vice-versa, in an insightful and practical way. In principle, this includes intervention strategies that use coalitions as a "stepping stone". But there are major technical problems to be overcome of two kinds: exactly how to generate specific model instances of sufficient complexity to be representative and informative, and how to interpret complex model instances once generated. Thus although there is substantial potential payoff, the prospect is a long-term and challenging one.

References

Abu-Zeid, M.A., and Biswas, A.K., (1996) "River Basin Planning and Management", Oxford University Press.

Bousquet F., Barreteau O., Le Page C., Mullon C. and Weber J. (1999) "An Environmental Modelling Approach: the Use of Multi-Agent Simulations", in "Advances in Environmental and Ecological Modelling" (F. Blasco and A. Weill, eds.), Paris, Elsevier, pp. 113-122.

Conte, R. and Gilbert, N. (1995) "Introduction: Computer Simulation for Social Theory", in "Artificial Societies: the Computer Simulation of Social Life" (N. Gilbert and R. Conte, eds.), UCL Press, London, pp. 1-15.

Doran, J. E. (1997) "From Computer Simulation to Artificial Societies", Transactions of the Society for Computer Simulation International, 14, 69-78

Doran, J. E. (2001) "Intervening to Achieve Co-operative Ecosystem Management: Towards an Agent Based Model", Journal of Artificial Societies and Social Simulation, 4(2), 2001, ON LINE: http://www.soc.surrey.ac.uk/JASSS/4/2/4.html

Doran, J. E. (2001a) "Can agent-based modelling REALLY be useful?", in "Cooperative Agents: Applications in the Social Sciences" (N. J. Saam and B. Schmidt, eds.), Philosophy and Methodology of the Social Sciences Series, Volume 32, Kluwer Academic Publishers, Dordrecht.

Doran, J. E., Franklin, S., Jennings, N. R. and Norman T. J. (1997) "On Cooperation in Multi-Agent Systems", The Knowledge Engineering Review, 12(3), 309-314.

Dorcey, A. H. J. (1997) "Collaborating Towards Sustainability Together: The Fraser Basin Management Board and Program", in "Practising Sustainable Water Management: Canadian and International Experiences" (D. Shrubsole and B. Mitchell, eds.), Canadian Water Resources Association. http://www.interchg.ubc.ca/dorcey/chcwra/fccwra.html

Epstein, J. M. and Axtell, R. (1996) "Growing Artificial Societies: Social Science from the Bottom Up", The Brookings Institution Press, Washington, D.C., The MIT Press Cambridge, Mass.

Gilbert, G. N. (2000) "Modelling Sociality: The View from Europe", in "Dynamics in Human and Primate Societies" (T. A. Kohler and G. J. Gumerman, eds.), Oxford University Press, pp. 355-372.

Hardin, G. (1968) "The Tragedy of the Commons", Science, 162, 1243-1248

Healey, M., ed. (1999) "Seeking Sustainability in the Lower Fraser Basin: Issues and Choices", Institute for Resources and the Environment, Westwater Research Centre, University of British Columbia, Vancouver.

Keohane, R. O. and Ostrom, E. eds. (1995) "Local Commons and Global Interdependence: Heterogeneity and Cooperation in Two Domains", Sage Publications, London.

Marshall, D. (1998) "Watershed management in British Columbia: The Fraser Basin Experience", Environments, 25(2/3), 64-79.

Moss, S. and Davidsson, P. eds. (2001) "Multi-Agent-Based Simulation" Proceedings of the Second International Workshop on Multi-Agent Based Simulation (MABS 2000), Boston, MA, USA, 8-9th July, 2000, Springer, Berlin

Ostrom, E. (1990) "Governing the Commons: The Evolution of Institutions for Collective Action". Cambridge University Press: Cambridge, UK.

Ostrom, E. (1995) "Constituting Social Capital and Collective Action", in: "Local Commons and Global Interdependence: Heterogeneity and Cooperation in Two Domains" (R. O. Keohane and E. Ostrom, eds.), Sage Publications, London. pp. 125-160.

Prietula, M. J., Carley, K.M., and Gasser, L. (1998) "Simulating Organizations", AAAI Press/The MIT Press Menlo Park, California, Cambridge, Mass., and London, England.

Russell, S. and Norvig, P. (1995) "Artificial Intelligence: a Modern Approach", Prentice-Hall.

Shoham, Y. (1993) "Agent-Oriented Programming", Artificial Intelligence, 60(1), 51-92.

Tessier C., Chaudron L., Müller H-J., eds. (2000) "Conflicting Agents: Conflict Management in Multi-Agent Systems", Kluwer Academic Publishers, Dordrecht.

Turing, A.M. (1937) "On Computable Numbers with an Application to the Entscheidungs-problem", Proc. London Math. Soc., 42, 230-265; corrections, ibid, 43 (1937) 544-6.

Weiss, G. ed. (1999) "Multiagent Systems", The MIT Press, Cambridge, Massachusetts, and London, England.

Westervelt, J., (2000) "Simulation Modeling for Watershed Management", Springer-Verlag, Telos

Wooldridge, M. and Jennings N.R. (1999) "The Cooperative Problem Solving Process", Journal of Logic and Computation, 9(4), 563-592